On the Capacity of Interference Channels with Degraded Message sets

Wei Wu, Sriram Vishwanath and Ari Arapostathis

Abstract— **This paper is motivated by a sensor network on a correlated field where nearby sensors share information, and can thus assist rather than interfere with one another. A special class of two-user Gaussian interference channels (IFCs) is considered where one of the two transmitters knows both the messages to be conveyed to the two receivers (called the IFC with degraded message sets). Both achievability and converse arguments are provided for this scenario for a class of discrete memoryless channels with weak interference. For the case of the Gaussian weak interference channel with degraded message sets, optimality of Gaussian inputs is also shown, resulting in the capacity region of this channel.**

Index Terms— **Network information theory, Interference channel, Dirty-paper coding**
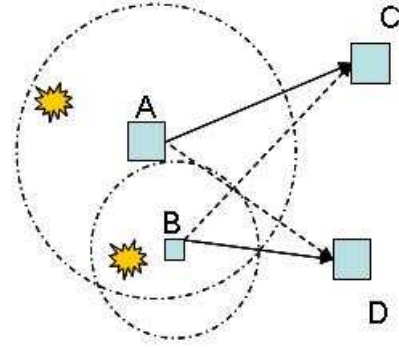
I. INTRODUCTION

An interference channel (IFC), characterized by the channel $p(y_1, y_2 | x_1, x_2)$, is one of the basic building blocks of many networks and is thus considered a fundamental problem in multi-user information theory. The capacity region of this channel remains an open problem, with some special cases such as the strong-interference case being solved. One of the fundamental difficulties faced while attacking the IFC capacity problem is that, unlike the broadcast channel, no transmit-side cooperation is possible.

Cooperation among transmitters can improve the rates achievable by employing joint encoding, i.e. [1] [2], [3]. As an example of this cooperation, Maric, Yates and Kramer consider the capacity of the strong interference channel with common information [4] and unidirectional cooperation [5].

In this work, we consider a two-user IFC where we allow limited cooperation between the transmitters by means of permitting one of the transmitters to possess the message of the other, or in other words, the message sets are *degraded*. Such a cooperative IFC is of interest on its own merit. One example of this is shown in Figure 1(a), in which Sensor A has better sensing capability thus can detect both events while Sensor B can only detect one of them and the data need to be sent to different destinations. In this example, Sensor A can potentially improve Sensor B's transmission with the additional transmitter side information by cooperation. Another example is shown in Figure 1(b): W_1, W_2 is transmitted from A to D, E respectively; if the intermediate nodes B, C employ a decode-forward coding scheme with an inbuilt degraded broadcast channel, then B, C, D, E form an IFC

The authors are with Wireless Networking and Communications Group, Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712, USA (e-mail: {wwu,sriram,ari}@ece.utexas.edu).

with degraded message sets, where B can cooperate with C in transmitting its message.

(a) Data collection through IFC
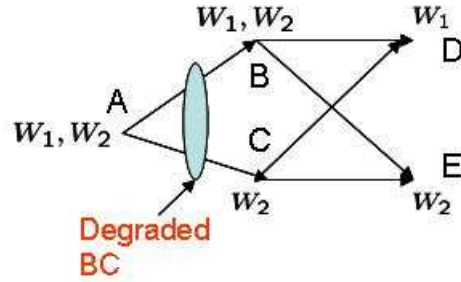
(b) Cascaded with degraded broadcast channel

Fig. 1. Two applications of interference channel with degraded message sets

In this paper, we find an achievable region and outer bound on the capacity region of the discrete memoryless IFCs with degraded message sets. Specifically, for a class of weak interference IFCs, this achievable region meets the outer bound giving us the capacity region for this class of channels. In the specific case of the Gaussian input distribution, we find Gaussian inputs to be optimal, resulting in the region being characterized in closed form.

The rest of the paper is organized as follows. In Section II, the basic definitions are presented and three types of weak interference are introduced. The main results are presented in Section III, including the capacity region of a class of

discrete memoryless weak interference channels (and Gaussian weak interference channels) with degraded message sets with numerical results of the Gaussian case shown in Section IV. Detailed proofs for Section III are given in Section V. Finally we conclude the paper with Section VI.

II. NOTATIONS AND PRELIMINARIES

A. Channel model and definitions

We adopt the following notational conventions. Random variables (RVs) will be noted by capital letters, while their realizations will be denoted by the respective lower case letters. X_m^n denotes the random vector $(X_m, \ldots, X_n)$, and X^n denotes the random vector $(X_1, \ldots, X_n)$.
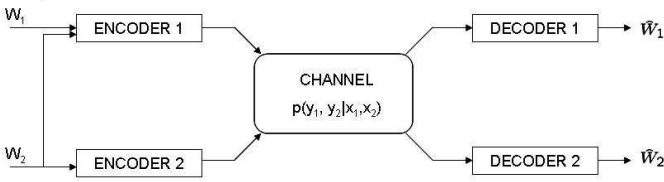

Fig. 2. The model of interference channel with degraded message sets

A two-user *interference channel* (IFC) $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2, P(y_1, y_2|x_1, x_2))$ is a channel with two input alphabets $\mathcal{X}_1$, $\mathcal{X}_2$, output alphabets $\mathcal{Y}_1$, $\mathcal{Y}_2$ and transition probability $P(y_1, y_2|x_1, x_2)$. It is assumed the channel is memoryless, namely

$$P(y_1^n, y_2^n|x_1^n, x_2^n) = \prod_{i=1}^{n} P(y_{1,i}, y_{2,i}|x_{1,i}, x_{2,i}).$$

Transmitter t sends a message with M_t bits, W_t, to receiver t in n channel uses at rate $R_t = M_t/n$ bits per use. A $(R_1, R_2, n, P_{e,1}, P_{e,2})$ code is defined as any code achieving the rate pair (R_1, R_2) with block size n and decoding error probability $P_{e,t}^{(n)}$, $t = 1, 2$. The *capacity region* $\mathcal{C}_{IFC}$ is the closure of the set of rate pairs (R_1, R_2), for which the receivers can decode their messages with error probability $P_{e,t}^{(n)} \to 0$ for $t = 1, 2$ as the block size $n \to +\infty$.

In the classic IFC described above, each transmitter has its own message set $\mathcal{M}_t = \{W_t\}$, where $W_t \in \{1, 2, \ldots, 2^{NR_t}\}$ denotes the private message to the receiver t. In the IFC with common information, which is proposed recently by Maric, Yates and Kramer in [4], each transmitter has not only its own private message W_t, but also the common message W_0 shared by all the transmitters. Thus the message set for transmitter t is $\mathcal{M} = \{W_0, W_t\}$.

In this paper, we consider the interference channel *with degraded message sets* (IFC-DMS). For a two-user IFC with the degraded message sets, the message set of one transmitter is a strict subset of the other. For example, Figure 2 corresponds to the message sets,

$$\{W_2\} = \mathcal{M}_2 \subset \mathcal{M}_1 = \{W_1, W_2\}, \tag{1}$$

for which the capacity region is denoted as $\mathcal{C}_{IFC}^{T_1}$ regarding to that transmitter 1 knows both messages. On the other hand,

$\mathcal{C}_{IFC}^{T_2}$ denotes the capacity region of IFC with transmitter 2 knowing both messages, namely

$$\{W_1\} = \mathcal{M}_1 \subset \mathcal{M}_2 = \{W_1, W_2\}. \tag{2}$$

In recent work (e.g. [5]), the IFC-DMS has also been referred to as *interference channel with unidirectional co-operation* - one transmitter knows the other's message and thus can enhance the achievable rate region. The definition of degradedness of message sets can be further generalized to K-user IFC: the message sets are degraded if there exists a permutation $\{\sigma_k, k = 1, \ldots, K\}$ of $\{1, 2, \ldots, K\}$, such that

$$\mathcal{M}_{\sigma_1} \subset \ldots \subset \mathcal{M}_{\sigma_K}.$$

In general, the capacity region of an IFC is an open problem and only known for certain classes of IFCs which include the so-called *strong interference* channels, which satisfy

$$\begin{aligned} I(X_1; Y_1|X_2) &\leq I(X_1; Y_2|X_2) \\ I(X_2; Y_2|X_1) &\leq I(X_2; Y_1|X_1), \end{aligned} \tag{3}$$

for all product distributions on the inputs X_1 and X_2. The capacity region in this case coinsides with the capacity region of compound IFC which is the union of two compound multiple access channels (MACs) determined by Ahlswede [6]. Under a strong interference assumption that is (slightly) different from (3), given by

$$\begin{aligned} I(X_1; Y_1|X_2, U) &\leq I(X_1; Y_2|X_2, U) \\ I(X_2; Y_2|X_1, U) &\leq I(X_2; Y_1|X_1, U), \end{aligned} \tag{4}$$

for all joint distributions $P(u, x_1, x_2, y_1, y_2)$ that factor as $P(u)P(x_1|u)P(x_2|u)P(y_1, y_2|x_1, x_2)$, Maric, Yates and Kramer find the capacity region of strong IFCs with common information [4] and with degraded message sets [5].

In this paper, we study three different notions of *weak interference* and investigate the capacity region for this channel with degraded message sets. These three notions are extensions of the concepts of *stochastically degraded*, *less noisy* and *more capable* from broadcast channel literature [7].

Definition 2.1: The interference channel is said to be *type A weak interference channel* if

$$I(X_1; Y_2|X_2) \leq I(X_1; Y_1|X_2) \tag{5}$$

or,

$$I(X_2; Y_1|X_1) \leq I(X_2; Y_2|X_1). \tag{6}$$

Definition 2.2: The interference channel is said to be *type B weak interference channel* if, for any U, $U \to (X_1, X_2) \to (Y_1, Y_2)$ forms a Markov chain, it satisfies

$$I(U, X_2; Y_2) \leq I(U, X_2; Y_1) \tag{7}$$

or,

$$I(U, X_1; Y_1) \leq I(U, X_1; Y_2). \tag{8}$$

Definition 2.3: The interference channel is said to be *type C weak interference channel* if there exists a probability transition matrix $q_1(y_2|x_2, y_1)$ such that

$$p(y_2|x_1, x_2) = \sum_{y_1} p(y_1|x_1, x_2)q_1(y_2|x_2, y_1), \tag{9}$$

or if there exists a probability transition matrix $q_2(y_1|x_1, y_2)$ such that

$$p(y_1|x_1, x_2) = \sum_{y_1} p(y_2|x_1, x_2)q_2(y_1|x_1, y_2). \quad (10)$$

Note that if the channel input X_2 of IFC is fixed, then the channel between X_1, Y_1, Y_2 forms a broadcast channel. For type C weak interference satisfying (9), this broadcast channel is *stochastically degraded*. This broadcast channel is *less noisy* for type B weak interference satisfying (7), and *more capable* for type A weak interference satisfying (5) [7]. Note that the notion of "more capable" is strictly weaker than that of "less noisy", which is strictly weaker than "degraded".

In this paper, we will establish the capacity region of the interference channel with degraded message sets in Figure 2 and type C weak interference.

B. Gaussian interference channels

One of our main interests in this paper is the Gaussian IFC, in which the alphabets of inputs and outputs are real numbers and the outputs are linear combinations of input signals and white Gaussian noise. The Gaussian IFC is defined as follows,

$$\begin{aligned} Y_1 &= X_1 + aX_2 + Z_1 \\ Y_2 &= bX_1 + X_2 + Z_2 \end{aligned} \quad (11)$$

where a and b are real numbers and Z_1 and Z_2 are independent, zero-mean, unit-variance Gaussian random variables. Furthermore, the transmitters are subject to average power constraints:

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[X_{tn}^2] \leq P_t, \quad t = 1, 2. \quad (12)$$

The capacity region of the non-cooperative Gaussian IFC is currently characterized for the cases when $a = b = 0$ (trivial) or if $a^2 \geq 1$ and $b^2 \geq 1$, in which case the strong interference conditions in (3) are satisfied. The capacity of IFC for strong interference is the set of (R_1, R_2) satisfying [8], [9]

$$0 \leq R_1 \leq \frac{1}{2} \log(1 + P_1) \quad (13)$$

$$0 \leq R_2 \leq \frac{1}{2} \log(1 + P_2) \quad (14)$$

$$0 \leq R_1 + R_2 \leq \frac{1}{2} \log(P_1 + a^2 P_2 + 1) \quad (15)$$

$$0 \leq R_1 + R_2 \leq \frac{1}{2} \log(b^2 P_1 + P_2 + 1). \quad (16)$$

In the regime when $0 \leq a^2 \leq 1$ or $0 \leq b^2 \leq 1$, the Gaussian IFCs belong to all three notions of weak interference. Achievable rate regions [9]–[11] and outer bounds [10], [12]–[14] are known for this scenario, but a characterization of the region is yet to be obtained. A recent outer bound by Kramer in [14] is given by (R_1, R_2) satisfying (13), (14), and
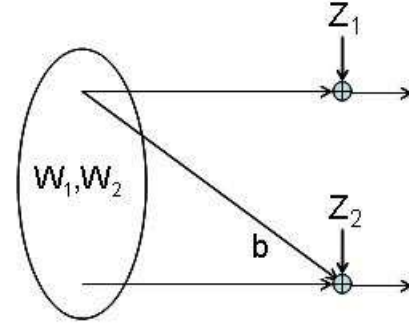
$$R_1 + R_2 \leq \frac{1}{2} \log\left[(P_1 + a^2 P_2 + 1)\left(\frac{P_2 + 1}{\min(a^2, 1)P_2 + 1}\right)\right] \quad (17)$$

$$R_1 + R_2 \leq \frac{1}{2} \log\left[(P_2 + b^2 P_1 + 1)\left(\frac{P_1 + 1}{\min(b^2, 1)P_1 + 1}\right)\right]. \quad (18)$$

Let the capacity region of a Gaussian IFC with degraded message sets (i.e., Figure 3(a)) be denoted by $\mathcal{C}_{GIFC}^{T_1}$. As a first step, we present an outer bound for $\mathcal{C}_{GIFC}^{T_1}$.
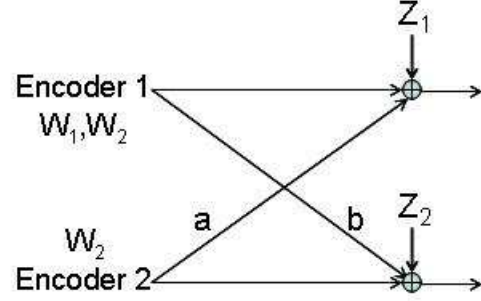
(a) The Gaussian IFC with degraded message sets

(b) The Gaussian MIMO broadcast channel

Fig. 3. The Gaussian MIMO BC outer bound for IFC-DMS

C. Outer bound for Gaussian IFC-DMS

First, we provide an "intuitive framework" for an outer bound on this channel. A formal outer bound follows this framework.

A straightforward outer bound for Gaussian IFC with degraded message sets (see Figure 3(a)), $\mathcal{C}_{GIFC}^{T_1}$, is the capacity region of the Gaussian broadcast channel resulting from allowing full transmitter-side cooperation. One can make an even tighter outer bound by using the following arguments:

  i) Removing the interference link from encoder 2 to receiver 1 in Figure 3(a) does not enhance the overall capacity region because encoder 2 does not have any knowledge of message W_1 and thus no cooperation can be induced to improve the transmission rate R_1;

 ii) Now allowing full cooperation between the two encoders provides us with an new broadcast channel shown in Figure 3(b), which has two transmit antennas and one receive antenna at each receiver and an individual power constraint at each antenna.

Using the existing literature on the capacity region of Gaussian multi-antenna (MIMO) BC channel [15]; dirty paper coding (DPC) [16] optimizes this outer bound [17] [18]. If, in the dirty paper coding strategy, W_1 is encoded first and W_2 second, the rates achieved are given by:

$$\mathcal{R}_{DPC}^{12} = \Big\{ (R_1, R_2): \quad \text{for} \quad 0 \le \alpha \le 1,$$
$$R_1 \le \frac{1}{2} \log \big(1 + \alpha P_1 \big),$$
$$R_2 \le \frac{1}{2} \log \big(\frac{1 + P_1 b^2 + 2|b| \sqrt{(1-\alpha) P_1 P_2} + P_2}{1 + b^2 \alpha P_1} \big) \Big\}. \quad (19)$$

The DPC achievable region for the encoding sequence W_2, W_1 is

$$\mathcal{R}_{DPC}^{21} = \Big\{ (R_1, R_2): \quad \text{for} \quad 0 \le \alpha \le 1,$$
$$R_1 \le \frac{1}{2} \log \big(\frac{1 + P_1}{1 + (1-\alpha) P_1} \big),$$
$$R_2 \le \frac{1}{2} \log \big(1 + P_1 b^2 + 2|b| \sqrt{(1-\alpha) P_1 P_2} + P_2 \big) \Big\}. \quad (20)$$

It is not hard to see that the capacity region of Gaussian IFC with degraded message set in Figure 3(a) is contained in the union of these two regions

$$\mathcal{C}_{GIFC}^{T_1} \subset \mathcal{R}_{DPC}^{12} \cup \mathcal{R}_{DPC}^{21} = \mathcal{R}_{DPC}^{21}.$$

We show later in this paper that the capacity region $\mathcal{C}_{GIFC}^{T_1}$ under weak interference is indeed equal to $\mathcal{R}_{DPC}^{12}$, and thus, even though the outer bound above is in general loose, it captures the intuition behind the optimal coding strategy for this channel.

III. MAIN RESULTS

In this section, we will first give our inner and outer bounds for the discrete-memoryless weak interference channel with degraded message sets, and then specialize to Gaussian weak IFC-DMS and establish its capacity region.

Define $\mathcal{R}_{in}$ to be the set of all rate (R_1, R_2) such that

$$R_1 \le I(V; Y_1) - I(V; U, X_2)$$
$$R_2 \le I(U, X_2; Y_2), \quad (21)$$

for the probability distribution $p(x_1, x_2, u, v, y_1, y_2)$ that factors as

$$p(u, x_2) p(v|u, x_2) p(x_2|v) p(y_1, y_2|x_1, x_2).$$

The following theorem gives the achievable region of IFC with transmitter 1 knowing both messages as in Figure 2 using the Gel'fand-Pinsker coding scheme [19].

Theorem 3.1: The capacity region of discrete memoryless interference channel with degraded message sets as (1) satisfies

$$\mathcal{R}_{in} \subset \mathcal{C}_{IFC}^{T_1}.$$

The proof is based on the Gel'fand-Pinsker coding scheme, in which (U, X_2) is considered as the random parameters for the channel between X_1 and Y_1. It is given in Section V-A.

The outer bound is stated next. Define $\mathcal{R}_o$ to be the set of all rate pairs (R_1, R_2) such that

$$R_1 \le I(X_1; Y_1 | X_2)$$
$$R_2 \le I(U, X_2; Y_2) \quad (22)$$
$$R_1 + R_2 \le I(X_1; Y_1 | U, X_2) + I(U, X_2; Y_2),$$

for the probability distribution $p(x_1, x_2, u, y_1, y_2)$ that factors as

$$p(u, x_2) p(x_1|u) p(y_1, y_2|x_1, x_2).$$

Theorem 3.2: The capacity region of discrete memoryless interference channel with degraded message sets satisfies

$$\mathcal{C}_{IFC}^{T_1} \subset \mathcal{R}_o.$$

The proof is provided in Section V-B. Both Theorem 3.1 and Theorem 3.2 hold for the general interference channels. However, as we can see, there is a gap between the achievable region obtained in Theorem 3.1 and the outer bound in Theorem 3.2.

Next we put additional weak interference assumptions (type B and type C) and consider its capacity region. The key idea is to find a rate region, which is the achievable region under type B weak interference condition and the outer bound under type C weak interference condition, thus the capacity region of type C weak interference channels.

Define the rate region $\mathcal{R}_*$ to be the set of all rate pairs (R_1, R_2) such that

$$R_1 \le I(X_1; Y_1 | U, X_2)$$
$$R_2 \le I(U, X_2; Y_2), \quad (23)$$

for the probability distribution $p(x_1, x_2, u, y_1, y_2)$ that factors as

$$p(u, x_2) p(x_1|u) p(y_1, y_2|x_1, x_2).$$

For type B weak interference channels, we have the following achievable region,

Theorem 3.3: The capacity region of discrete memoryless type B weak interference channel with degraded message sets and that (7) holds, satisfies

$$\mathcal{R}_* \subset \mathcal{C}_{IFC}^{T_1, B}.$$

The proof is based on the coding scheme for degraded broadcast channel and the sketch of the proof is provided in Section V-C for completeness. Note since type C weak interference channels satisfy (7) also, thus

$$\mathcal{R}_* \subset \mathcal{C}_{IFC}^{T_1, C} \quad (24)$$

holds.

For type C weak interference channels, we can also show $\mathcal{R}_*$ is also its outer bound.

Theorem 3.4: The capacity region of discrete memoryless type C weak interference channel with degraded message sets and that (9) holds, satisfies

$$\mathcal{C}_{IFC}^{T_1, C} = \mathcal{R}_*.$$

The proof is given in Section V-D.

Next we will specialize the weak interference channel to the Gaussian case. Since Gaussian weak interference channels satisfy type C weak interference condition, the capacity region

of Gaussian weak interference channel with degraded message sets can be established by Theorem 3.4 and proving optimality of Gaussian input.

Theorem 3.5: The capacity region of the Gaussian IFC with Gaussian inputs and transmitter T_1 knowing both messages, $\mathcal{C}^{T_1}_{GIFC}$, when $|b| \leq 1$, is the set of all rate pairs (R_1, R_2) such that, for $0 \leq \alpha \leq 1$,

$$R_1 \leq \frac{1}{2} \log \left(1 + \alpha P_1\right) \tag{25}$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{h \Sigma h^t}{1 + b^2 \alpha P_1}\right)$$

$$= \frac{1}{2} \log \left(\frac{1 + P_1 b^2 + 2|b|\sqrt{(1-\alpha)P_1 P_2} + P_2}{1 + b^2 \alpha P_1}\right) \tag{26}$$

Here h is the vector $[b \ 1]$, and Σ is a 2×2 covariance with diagonal elements equaling $(1-\alpha)P_1$ and P_2 respectively.

As we have seen, the capacity region established in Theorem 3.5 is exactly equal to the dirty paper region in (19), i.e.,

$$\mathcal{C}^{T_1}_{GIFC} = \mathcal{R}^{12}_{DPC} \,,$$

when the encoding sequence is W_1, W_2 for the Gaussian BC channel.

By swapping the parameters of two transmitters, the capacity region $\mathcal{C}^{T_2}_{GIFC}$ can be obtained as the following corollary.

Corollary 3.6: The capacity region of the Gaussian IFC with Gaussian inputs and transmitter 2 knowing both messages, $\mathcal{C}^{T_2}_{GIFC}$, when $|a| \leq 1$, is the set of all rate pairs (R_1, R_2) such that, for $0 \leq \beta \leq 1$,

$$R_1 \leq \frac{1}{2} \log \left(\frac{1 + P_1 + 2|a|\sqrt{(1-\beta)P_1 P_2} + a^2 P_2}{1 + a^2 \beta P_2}\right)$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \beta P_2\right) . \tag{27}$$

IV. NUMERICAL RESULTS

In this section, we compare the capacity region of Gaussian IFCs with degraded message sets with the outer bounds of the normal Gaussian IFCs and the rate regions of Gaussian BC channels by numerical results.

We consider the symmetric Gaussian interference channel with $P_1 = P_2 = 6$ and $a^2 = b^2 = 0.3$. The rate units are bits per channel use. First we compare the capacity region of GIFC-DMS, $\mathcal{C}^{T_1}_{GIFC}$ with the dirty paper coding regions $\mathcal{R}^{12}_{DPC}$ in (19), $\mathcal{R}^{21}_{DPC}$ in (20) and the outer bound of GIFC in (17) – (18). As we can see, the capacity region of GIFC-DMS is strictly larger than the outer bound of GIFC, and the gap between these two shows the performance improvement by allowing encoders to cooperate partially. The point when $\alpha = 0$ corresponds to the full cooperation between two encoders to transmit message W_2.

On the other hand, the capacity region of GIFC-DMS can be served as the outer bound for GIFC. Since both $\mathcal{C}^{T_1}_{GIFC}$ and $\mathcal{C}^{T_2}_{GIFC}$ include the capacity region of GIFC, we have

$$\mathcal{C}_{GIFC} \subset \mathcal{C}^{T_1}_{GIFC} \cap \mathcal{C}^{T_1}_{GIFC} \,. \tag{28}$$
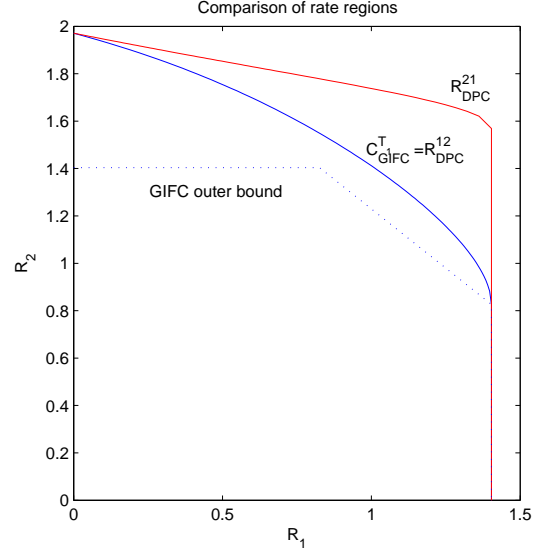

Fig. 4. The capacity region of Gaussian IFC-DMS with $P_1 = P_2 = 6$, $a^2 = b^2 = 0.3$ and achievable rate regions of Gaussian BC and IFC

In Figure 5, we compare some known outer bounds for two-user Gaussian IFC for the same setup. As we can see, Kramer's outer bound in [14] meets our outer bound of (28) at point A, B, and performs better than ours elsewhere. Our outer bound in (28) gives a better bound than Carleial's when α and β in (25) (26) (27) is close to 1.

Note the outer bounds (17), (18) are proved based on a genie-aided argument, or more precisely "receiver genie-aided", in which some genie provides additional channel output to one of receivers. In contrast, our approach here can be viewed as "transmitter genie-aided" in the sense that some genie gives additional information of channel input to one of the transmitters. We comment that the outer bound by (28) is not as good as the one obtained in [14]. However, the bound in (28) might be possibly improved when one transmitter only knows partial information of the other message instead of the full message: for example, in Figure 2 transmitter T_1 only know $g(W_2)$ instead of knowing W_2, where $g(\cdot)$ is some function of W_2.

V. PROOFS

A. Proof of Theorem 3.1

The coding scheme is a straightforward application of Gelfand-Pinsker coding.

Code Generation: Fix $p(u, x_2)$, first generate 2^{nR_2} independent codewords of length n at random according to the distribution $\prod_{i=1}^{n} p(u_i, x_{2,i})$ for message $w_2 \in \{1, \ldots, 2^{nR_2}\}$. Then generate $2^{nI(V;Y)} - \epsilon$ i.i.d. sequences v^n according to the distribution $\prod_{i=1}^{n} p(v_i)$, and distribute these sequences uniformly into 2^{nR_1} bins. For each sequence v^n, let $i(v^n)$ denote the index of the bin containing v^n.

Encoding: Encoder 2 transmits $X_2(w_2)$. With w_1 and w_2, Encoder 1 looks in bin w_1 for a sequence V^n that $(X_1^n, V^n(w_1), X_2^n(w_2), U^n(w_2))$ is jointly typical and sends
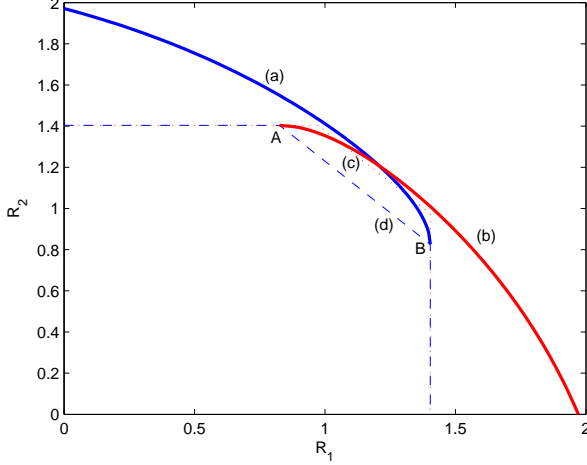
Fig. 5. The outer bounds of two-user Gaussian interference channel with $P_1 = P_2 = 6$, $a^2 = b^2 = 0.3$: (a) the capacity region $\mathcal{C}_{IFC}^{T_1}$; (b) the capacity region $\mathcal{C}_{IFC}^{T_2}$; (c) Carleial's outer bound in [13]; (d) Kramer's outer bound in [14] (theorem 1).

$X_{1,i}$. If no such a sequence, an error is declared. If the number of sequences in each bin is larger than $2^{I(U,X_2;V)}$, the probability of finding no such U^n decreases to 0 as n goes to $+\infty$.

Decoding: Receiver 2 determines the unique $\hat{\hat{W}}_2$ such that $(U^n(\hat{\hat{W}}_2), X_2^n(\hat{\hat{W}}_2), Y_2^n)$ is jointly typical. Receiver 1 looks for the unique V^n such that (V^n, Y_1^n) is jointly typical and estimate the message $\hat{W}_1$ as the index of the bin containing the obtained V^n.

At Receiver 1, if

$$R_1 \leq I(V; Y_1) - I(V; U, X_2) - \epsilon,$$

the probability of error of w_1 decreases exponentially to zero as $n \to \infty$. The probability of error of w_2 at receiver 2 goes to zero as $n \to +\infty$ if $R_2 \leq I(U, X_2; Y_2)$.

B. Proof of Theorem 3.2

Theorem 3.2 can be proved by adapting Marton's BC outer bound. For a $(R_1, R_2, n, P_{e,1}^{(n)}, P_{e,2}^{(n)})$ code with decoding error $P_{e,i}^{(n)} \to 0$ as $n \to +\infty$, we define the auxiliary random variable U,

$$U_i = (W_2, Y_1^{i-1}, Y_{2,i+1}^n, X_2^{i-1}, X_{2,i+1}^n). \quad (29)$$

Applying Fano's inequality [20] for each message W_t, $t = 1, 2$, we have

$$H(W_t|Y_t^n) \leq nR_t P_{e,t}^{(n)} + h(P_{e,t}^{(n)}) = n\epsilon_t^{(n)}, \quad (30)$$

where $\epsilon_t^{(n)} \to 0$ as $P_{e,t}^{(n)} \to 0$. Moreover, because transmitter 2 has no information about the message W_1, X_2^n is independent of W_1 or the following relation holds

$$H(W_1|W_2, X_2^n) = H(W_1). \quad (31)$$

To prove the converse, we need following lemmas:

Lemma 5.1 ([21]): For any random variable T, the following equality holds,

$$\sum_{i=1}^n I(Y_{2,i+1}^n; Y_{1,i}|Y_1^{i-1}, T)$$

$$= \sum_{i=1}^n I(Y_1^{i-1}; Y_{2,i}|Y_{2,i+1}^n, T). \quad (32)$$

First let prove the outer bound for R_1 in (22).

$$nR_1 \leq I(W_1; Y_1^n|X_2^n) + n\epsilon_1^{(n)} \quad (33)$$

$$= \sum_{i=1}^n I(W_1; Y_{1,i}|Y_1^{i-1}, X_2^n) + n\epsilon_1^{(n)}$$

$$\leq \sum_{i=1}^n \Big(H(Y_{1,i}|X_{2,i}) - H(Y_{1,i}|X_2^n, Y_1^{i-1}, X_{1,i}) \Big)$$

$$+ n\epsilon_1^{(n)}$$

$$\leq \sum_{i=1}^n I(X_{1,i}; Y_{1,i}|X_{2,i}) + n\epsilon_1^{(n)}.$$

Next we prove the outer bound for R_2 in (22):

$$nR_2 \leq I(W_2; Y_2^n) + n\epsilon_2^{(n)} \quad (34)$$

$$\leq \sum_{i=1}^n I(W_2; Y_{2,i}|Y_{2,i+1}^n) + n\epsilon_2^{(n)}$$

$$\leq \sum_{i=1}^n I(W_2, Y_{2,i+1}^n; Y_{2,i}) + n\epsilon_2^{(n)}$$

$$\leq \sum_{i=1}^n I(U_i, X_{2,i}; Y_{2,i}) + n\epsilon_2^{(n)},$$

where (33) and (34) are from Fano's inequality in (30)

Then we prove the outer bound for the sum rate $R_1 + R_2$ in (22):

$$n(R_1 + R_2)$$

$$\leq I(W_1; Y_1^n|W_2, X_2^n) + I(W_2; Y_2^n) + n\epsilon_1^{(n)} + n\epsilon_2^{(n)} \quad (35a)$$

$$\leq \sum_{i=1}^n \Big(I(W_1; Y_{1,i}|W_2, X_2^n, Y_1^{i-1}) + I(W_2; Y_{2,i}|Y_{2,i+1}^n) \Big)$$

$$+ n\epsilon_1^{(n)} + n\epsilon_2^{(n)}$$

$$\leq \sum_{i=1}^n \Big(I(W_1; Y_{1,i}|W_2, Y_1^{i-1}, Y_{2,i+1}^n, X_2^n)$$

$$+ I(Y_{2,i+1}^n; Y_{1,i}|W_2, Y_1^{i-1}, X_2^n)$$

$$+ I(W_2, Y_1^{i-1}, Y_{2,i+1}^n, X_2^n; Y_{2,i})$$

$$+ I(Y_1^{i-1}; Y_{2,i}|W_2, Y_{2,i+1}^n, X_2^n) + \epsilon_1^{(n)} + \epsilon_2^{(n)} \Big) \quad (35b)$$

$$\leq \sum_{i=1}^n \Big(I(W_1; Y_{1,i}|U_i, X_{2,i}) + I(U_i, X_{2,i}; Y_{2,i}) \Big)$$

$$+ n\epsilon_1^{(n)} + n\epsilon_2^{(n)} \quad (35c)$$

$$\leq \sum_{i=1}^n \Big(I(X_1; Y_{1,i}|U_i, X_{2,i}) + I(U_i, X_{2,i}; Y_{2,i}) \Big)$$

$$+ n\epsilon_1^{(n)} + n\epsilon_2^{(n)}. \quad (35d)$$

Note (35a) is due to Fano's inequality in (30) and the conditional entropy relation in (31); the two terms in (35b) are equal due to Lemma 5.1; (35c) is obtained by using the auxiliary random variable U_i and (35d) is true because $(W_1, U_i) \rightarrow (X_{1,i}, X_{2,i}) \rightarrow (Y_{1,i}, Y_{2,i})$ forms a Markov chain in this order for all $1 \leq i \leq n$.

C. Proof of Theorem 3.3

Here we provide a proof for Theorem 3.3, the achievability of $\mathcal{R}_*$ for type B weak interference.

Code Generation: Fix $p(u, x_2)$, generate 2^{nR_2} independent codewords of length n at random according to the distribution $\prod_{i=1}^{n} p(u_i, x_{2,i})$ for message $w_2 \in \{1, \ldots, 2^{nR_2}\}$. For each codeword $(U^n(w_2), X_2^n(w_2))$, generate 2^{nR_1} independent codewords $X(w_1, w_2)$ according to $\prod_{i=1}^{n} p(x_1|u, x_2)$, where $w_1 \in \{1, \ldots, 2^{nR_1}\}$.

Encoding: Encoder 2 transmits $X_2(w_2)$. Since encoder 1 knows both messages, it sends $X_1(w_1, w_2)$.

Decoding: Receiver 2 determines the unique $\hat{W}_2$ such that $(U^n(\hat{W}_2), X_2^n(\hat{W}_2), Y_2)$ is jointly typical. Receiver 1 looks for the unique $(\hat{W}_1, \hat{W}_2)$ such that $(X_1^n(\hat{W}_1, \hat{W}_2), X_2^n(\hat{W}_2), U^n(\hat{W}_2))$ is jointly typical.

It is easy to see that the probability of error at receiver 2 goes to zero as $n \rightarrow +\infty$ if $R_2 \leq I(U, X_2; Y_2)$. Receiver 1 can decode W_2 successfully as $n \rightarrow +\infty$ if $R_2 \leq I(U, X_2; Y_1)$. For type B weak interference with (7),

$$R_2 \leq I(UX_2; Y_2) \leq I(UX_2; Y_1).$$

Thus receiver 1 can decode W_2 as long as receiver 2 can do so. With the error probability of W_2 going to zero as $n \rightarrow +\infty$, the error probability of W_1 at receiver 1 goes to zero if $R_1 \leq I(X_1; Y_1|X_2, U)$. The above analysis shows that both receivers can decode with total probability of error going to 0 if (23) is satisfied. Hence there exists a sequence of good codes with error probability going to 0.

D. Proof of Theorem 3.4

For the proof of Theorem 3.2, the key is to identify the auxiliary random variable and utilize the definition of type B weak interference.

For any $(R_1, R_2, n, P_{e,1}^{(n)}, P_{e,2}^{(n)})$ code with decoding error $P_{e,i}^{(n)} \rightarrow 0$ as $n \rightarrow +\infty$, to prove the outer bound in (23) for R_1, we consider

$$nR_1 = H(W_1)$$
$$\leq I(W_1; Y_1^n) - n\epsilon_1^{(n)} \tag{36a}$$
$$\leq I(W_1; Y_1^n|W_2, X_2^n) - n\epsilon_1^{(n)} \tag{36b}$$
$$\leq \sum_{i=1}^{n} I(W_1; Y_{1,i}|W_2, Y_1^{i-1}, X_2^n) - n\epsilon_1^{(n)} \tag{36c}$$

where the inequality (36a) is due to Fano's inequality (30); (36b) follows from (31) and (36c) is due to the reason that the mutual information will be increases by adding conditionals.

Define $U_i = (W_2, Y_1^{i-1}, X_2^{i-1})$,

$$nR_1 \leq \sum_{i=1}^{n} I(W_1; Y_{1,i}|U_i, X_{2,i}, X_{2,i+1}^n) + n\epsilon_1^{(n)}$$
$$= \sum_{i=1}^{n} \Big(H(Y_{1,i}|U_i, X_{2,i}, X_{2,i+1}^n)$$
$$\quad - H(Y_{1,i}|U_i, X_{2,i}, X_{2,i+1}^n, W_1) \Big) + n\epsilon_1^{(n)} \tag{36d}$$
$$\leq \sum_{i=1}^{n} \Big(H(Y_{1,i}|U_i, X_{2,i})$$
$$\quad - H(Y_{1,i}|U_i, X_{2,i}, X_{2,i+1}^n, W_1, X_{1,i}) + \epsilon_1^{(n)} \Big) \tag{36e}$$
$$= \sum_{i=1}^{n} \Big(H(Y_{1,i}|U_i, X_{2,i})$$
$$\quad - H(Y_{1,i}|U_i, X_{1,i}, X_{2,i}) + \epsilon_1^{(n)} \Big) \tag{36f}$$
$$= \sum_{i=1}^{n} \Big(I(X_{1,i}; Y_{1,i}|U_i, X_{2,i}) + \epsilon_1^{(n)} \Big) \tag{36g}$$

(36e) is because the entropy increases when dropping some conditionals and it decreases by adding more conditions. The equality (36f) is true because $W_1 - (X_{1,i}, X_{2,i}) - Y_{1,i}$ forms a Markov chain.

Now for the bound of R_2 in (23), we have

$$nR_2 = H(W_2)$$
$$\leq I(W_2; Y_2^n) + n\epsilon_2^{(n)} \tag{37a}$$
$$= \sum_{i=1}^{n} I(W_2; Y_{2,i}|Y_2^{i-1}) + n\epsilon_2^{(n)}$$
$$\leq \sum_{i=1}^{n} \Big(H(Y_{2,i}) - H(Y_{2,i}|Y_2^{i-1}, W_2, X_2^i) + \epsilon_2^{(n)} \Big), \tag{37b}$$

where (37a) follows from Fano's inequality in (30); (37b) is because the entropy will decrease when adding conditionals.

Note for the interference channel two receivers cannot cooperate, the capacity region is the same as the one with the same marginal output $p(y_1|x_1, x_2)$, $p(y_2|x_1, x_2)$. Thus for type B weak interference satisfying (7), given X_2, one can construct a Markov chain, namely,

$$X_1 \rightarrow Y_1 \rightarrow Y_2.$$

Thus the conditional entropy in (37b) satisfies

$$H(Y_{2,i}|Y_2^{i-1}, W_2, X_2^i)$$
$$\geq H(Y_{2,i}|Y_1^{i-1}, W_2, X_2^i) \tag{38a}$$
$$= H(Y_{2,i}|U_i, X_{2,i}) \tag{38b}$$

Combine (24) and (37b), (38b), we get the result in Theorem 3.4.

E. Proof of Theorem 3.5

Achievability: The achievability of this rate utilizes the now famous dirty-paper coding strategy. First, we generate a codebook of 2^{nR_2} codewords according to $\mathcal{N}(0, \Sigma)$, where

Σ is the covariance between transmitter 1 and 2. Transmitter 1 devotes $(1-\alpha)$ fraction of its power P_1 to the transmission of W_2, while Transmitter 2 devotes its entire power P_2 to this effort. This leads to a covariance of the form

$$\Sigma = \begin{bmatrix} (1-\alpha)P_1 & \gamma \\ \gamma & P_2 \end{bmatrix} \tag{39}$$

The effective interference seen by Receiver 1 is a combination of the signals communicated from both Transmitters 1 and 2. Since Transmitter 1 knows the exact realization of the message $w_2 \in W_2$, it has non-causal side information on the interference and can completely cancel it out, achieving a rate

$$R_1 = \frac{1}{2}\log\left(1 + \alpha P_1\right) \tag{40}$$

using a Gaussian codebook with codewords that are correlated with the interference. At Receiver 2, this Gaussian codebook for W_1 is perceived as additive interference, hence achieving a rate:

$$R_2 = \frac{1}{2}\log\left(1 + \frac{h\Sigma h^t}{1 + b^2\alpha P_1}\right) \tag{41}$$

Maximize R_2 over $|\gamma|^2 \leq (1-\alpha)P_1 P_2$ (such that Σ is positive definite), it is not difficult to shown R_2 obtain the maximum when $\gamma = \sqrt{(1-\alpha)P_1 P_2}$ and (26) can be achieved.

Converse: Since Gaussian IFCs with degraded message sets and $|b| \leq 1$ satisfy the condition (9) for type C weak interference, the outer bound (23) holds. The central feature here is to prove the optimality of Gaussian input.

To prove the Gaussian optimality, we need following lemmas:

Lemma 5.2 (Lemma 1 in [22]): Let $X_1, X_2, \ldots, X_k$ be an arbitrary set of zero-mean random variables with covariance matrix K. Let S by any subset of $\{1, 2, \ldots, k\}$ and $\bar{S}$ be its complement. Then

$$h(X_S | X_{\bar{S}}) \leq h(X_S^* | X_{\bar{S}}^*),$$

where

$$(X_1^*, X_2^*, \ldots, X_k^*) \sim N(0, K).$$

Lemma 5.3: Let X_1, X_2 be arbitrarily distributed zero-mean random variables and X_1^*, X_2^* be zero-mean Gaussian distributed random variables with the same covariance matrix as X_1, X_2, then

$$\mathbb{E}[X_1 X_2] \leq \left(\mathbb{E}\left[(\mathbb{E}[X_1^*|X_2^*])^2\right] \mathbb{E}\left[(X_2^*)^2\right]\right)^{\frac{1}{2}}$$

Proof: Using Cauchy inequality and the properties of conditional expectation,

$$\begin{aligned} \mathbb{E}[X_1 X_2] &= \mathbb{E}[X_1^* X_2^*] \\ &= \mathbb{E}\left[\mathbb{E}[X_1^* X_2^* | X_2^*]\right] \\ &= \mathbb{E}\left[\mathbb{E}[X_1^* | X_2^*] X_2^*\right] \\ &\leq \left(\mathbb{E}\left[(\mathbb{E}[X_1^*|X_2^*])^2\right] \mathbb{E}\left[(X_2^*)^2\right]\right)^{\frac{1}{2}}. \end{aligned}$$

■

Let denote X_1^*, X_2^*, U^* are Gaussian distributed random variables with the same covariance with X_1, X_2, U, and

$Var(()A|B)$ denotes the conditional variance of A given B, namely,

$$Var(A|B) = \mathbb{E}[A^2|B].$$

First we note that

$$\begin{aligned} 0 &\leq h(Y_1|U, X_2) \\ &= h(X_1 + aX_2 + Z_1|U, X_2) \\ &= h(X_1 + Z_1|U, X_2) \\ &\leq h(X_1 + Z_1) \\ &\leq \frac{1}{2}\log(1 + P_1). \end{aligned}$$

Without losing any generality, we can assume

$$h(Y_1|U, X_2) = \frac{1}{2}\log(1 + \alpha P_1), \tag{42}$$

for some α, $0 \leq \alpha \leq 1$. Thus we have

$$\begin{aligned} I(X_1; Y_1|U, X_2) &= h(Y_1|U, X_2) - h(Y_1|U, X_1, X_2) \\ &= \frac{1}{2}\log(1 + \alpha P_1) \end{aligned} \tag{43}$$

On the other hand, by Lemma 5.2, the conditional entropy is upper bounded by the Gaussian variables with the same covariance matrix, thus,

$$\begin{aligned} h(Y_1|U, X_2) &\leq h(X_1 + Z_1|X_2) \\ &\leq h(X_1^* + Z_1|X_2^*) \\ &= \frac{1}{2}\log\left(1 + Var(X_1^*|X_2^*)\right). \end{aligned}$$

Combining with (42), we have a lower bound for $Var(X_1^*|X_2^*)$,

$$Var(X_1^*|X_2^*) \geq \alpha P_1. \tag{44}$$

Note

$$Var(X_1^*|X_2^*) = \mathbb{E}[(X_1^*)^2] - \mathbb{E}\left[(\mathbb{E}[X_1^*|X_2^*])^2\right],$$

together with (44), we have

$$\mathbb{E}\left[(\mathbb{E}[X_1^*|X_2^*])^2\right] \leq (1-\alpha)P_1. \tag{45}$$

Combine Lemma 5.3 and (45),

$$\mathbb{E}[X_1 X_2] \leq \sqrt{(1-\alpha)P_1 P_2}.$$

Thus

$$\begin{aligned} h(Y_2) &= h(bX_1 + X_2 + Z_2) \\ &\leq \frac{1}{2}\log(1 + b^2 P_1 + P_2 + 2b\,\mathbb{E}[X_1 X_2]) \\ &\leq \frac{1}{2}\log(1 + b^2 P_1 + P_2 + 2b\sqrt{(1-\alpha)P_1 P_2}). \end{aligned} \tag{46}$$

At last we need to bound $h(Y_2|X_2, U)$. Define Y_1', Y_2' as follows,

$$\begin{aligned} Y_1' &= X_1 + Z_1 \\ Y_2' &= bX_1 + Z_2. \end{aligned}$$

Thus Y_2' is a stochastically degraded version of Y_1', or in other words, $Y_2' = bY_1' + Z'$, where Z' is Gaussian distributed with variance $1 - b^2$. By entropy power inequality (EPI) [20],

$$2^{2h(Y_2|X_2,U)} = 2^{2h(Y_2'|X_2,U)}$$
$$\geq 2^{2h(bY_1'|X_2,U)} + 2^{2h(Z')}$$
$$= b^2 2^{2h(Y_1|X_2,U)} + 1 - b^2$$
$$= 1 + b^2\alpha P_1\,,$$

thus

$$h(Y_2|X_2,U) \geq \frac{1}{2}\log(1 + b^2\alpha P_1)\,. \tag{47}$$

And finally we combine (46) and (47),

$$I(X_2,U;Y_2)$$
$$= h(Y_2) - h(Y_2|X_2,U)$$
$$\leq \frac{1}{2}\log\Big(\frac{1 + b^2 P_1 + P_2 + 2b\sqrt{(1-\alpha)P_1 P_2}}{1 + b^2\alpha P_1}\Big)\,. \tag{48}$$

Since (43) (43) are similar to (25) (26), the optimality of Gaussian inputs is established and the proof for converse is complete.

VI. Conclusions and Future Work

In this paper, we investigate the capacity region of two-user interference channel (IFC) with degraded message sets (DMS), in which one transmitter knows both messages. For the general discrete memoryless interference channel setting, we find an achievable region and outer bound, which meet under a weak interference condition. Thus, we determine the capacity region of a class of channels that includes the Gaussian weak interference channel.

Possible extensions to this work include:

i) Lossy functions of the message (w_2) made available to T_1 rather than the message itself. This might possibly yield a better outer bound for the non-cooperative IFC.

ii) Generalizing this approach to IFC with more than two users in the system.

Acknowledgments

References

[1] U. M. N. Jindal and A. Goldsmith, "Capacity of ad-hoc networks with node cooperation," in *Proceedings of IEEE Int. Symp. Inf. Theory*, 2004, p. 271.

[2] A. Host-Madsen, "A new achievable rate for cooperative diversity based on generalized writing on dirty paper," in *Proceedings of IEEE Int. Symp. Inf. Theory*, June 2003, p. 317.

[3] C. Ng and A. Goldsmith, "Transmitter cooperation in ad-hoc wireless networks: Does dirty-paper coding beat relaying?" in *Proceedings of IEEE Information Theory Workshop*, Oct. 2004.

[4] I. Maric, R. D. Yates, and G. Kramer, "The strong interference channel with common information," in *Proceedings of 43th Allerton Conference on Communications, Control and Computing, Monticello, IL*, Sep. 2005.

[5] ——, "The strong interference channel with unidirectional cooperation," in *Proceedings of the UCSD Workshop on Information Theory and its Applications, San Diego, CA*, Feb. 2006.

[6] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Annals of Probability*, vol. 2, no. 5, pp. 805–814, 1974.

[7] A. E. Gamal, "The capacity of a class of broadcast channels," *IEEE transactions on information theory*, vol. 25, no. 2, pp. 166–169, Mar. 1979.

[8] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE transactions on information theory*, vol. 27, pp. 786–788, Nov. 1981.

[9] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE transactions on information theory*, vol. 27, pp. 49–60, Jan. 1981.

[10] M. H. Costa, "On the Gaussian interference channel," *IEEE transactions on information theory*, vol. 31, pp. 607–615, Sep. 1985.

[11] I. Sason, "On achievable rate regions for the Gaussian interference channel," *IEEE transactions on information theory*, vol. 50, pp. 1345–1356, Jun. 2004.

[12] H. Sato, "Two-user communication channels," *IEEE transactions on information theory*, vol. 23, pp. 295–304, May 1977.

[13] A. Carleial, "Outer bounds on the capacity of interference channels," *IEEE transactions on information theory*, vol. 29, pp. 602–606, Jul. 1983.

[14] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE transactions on information theory*, vol. 50, pp. 581–586, Mar. 2004.

[15] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," *accepted by IEEE transactions on information theory*, 2005.

[16] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. on Information Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[17] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. on Information Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[18] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussin broadcast channel and uplink-downlink duality," *IEEE Trans. Information Theory*, vol. 49, pp. 1912–1921, Aug. 2003.

[19] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Probl. Peredachi Inform. (Probl. Inform. Trans.)*, vol. 9, no. 1, pp. 19–31, 1980.

[20] T. M. Cover and J. A. Thomas, *Elements of information theory*, ser. Wiley Series in Telecommunications. New York: John Wiley & Sons Inc., 1991, a Wiley-Interscience Publication.

[21] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Budapest: Akademiai Kiado, 1997.

[22] J. A. Thomas, "Feedback can at most double Gaussian multiple access channel capacity," *IEEE transactions on information theory*, vol. 33, pp. 711–716, Sep. 1987.